MDDAC

PRACTITIONER GUIDELINES

Rachel Allsopp, Claire Bessant, Sneha Dawda, Keith Ditcham, Charlotte Emmett, Matthew Higgs, Ardi Janjeva, Guangquan Li, Selina Sutton and Mark Warner

Principal Investigator: Marion Oswald

UKRI Arts and Humanities Research Council

Northumbria University NEWCASTLE

RUSI www.rusi.org

INTRODUCTION

These practitioner guidelines are presented by the AHRC-funded 'Observatory for Monitoring Data-Driven Approaches to COVID-19' (OMDDAC) project. OMDDAC is a collaboration between Northumbria University and the Royal United Services Institute (RUSI), researching the data-driven approaches to COVID-19, with a focus upon legal, ethical, policy and operational challenges.

OMDDAC has analysed key data-driven responses to COVID-19, collating lessons learned in 'real-time' throughout the pandemic by way of representative public surveys, case study analysis and interviews with key stakeholders from a range of sectors (including local and central government, regulators, law enforcement, the medical and legal profession, charities and the third sector, the private sector, and an interdisciplinary range of academics). These practitioner guidelines have been informed by our research findings.

The guidelines are relevant specifically to practitioners who work with data in the health and social care sector and in the law enforcement sector.

The guidelines are underpinned by broad ethical standards which build on the Seven Principles of Public Life, more commonly known as the Nolan Principles. The principles of **selflessness, integrity, objectivity, accountability, openness, honesty** and **leadership;** plus **legality** (an eighth principle added by OMDDAC) are all relevant to the application of data-driven approaches and have been adapted from their initial meaning to reflect this. It is not intended that these standards serve as a technical framework; rather they provide a high-level specification against which practitioners can measure their actions over the course of a data-driven project or interaction, complementing our more detailed guidelines.

Selflessness: Taking the requisite steps to share data with the right people or bodies even if it appears arduous, and taking the initiative to do this proactively as well as on request.

Integrity: Ensuring that the objectives of data-driven policy are defined to serve the public good and not private interests.

Objectivity: Teams working collaboratively on data-driven outcomes should be made up of individuals from a wide range of backgrounds to improve the chances of impartiality and reach answers which minimise the likelihood of bias and discrimination.

Accountability, scrutinising data-driven decisions to the degree that the public have trust and confidence in them.

Openness: Making decisions in a transparent manner, publishing the underlying rationale and explanatory notices to ensure public understanding and acceptance. Not withholding information from the public without a clearly explained reason.

Honesty: Being clear about where things have gone (or may go) wrong, and where there are levels of uncertainty.

Leadership: Where these broad principles are not followed, senior management figures must challenge and be seen to actively remedy any failures.

Legality: Ensuring that any data use and response measures imposed are in accordance with the law, necessary and proportionate in order to achieve a legitimate aim, non-discriminatory and capable of challenge.

HEALTH AND SOCIAL CARE GUIDELINES

Target audience:

Policy and Decision-Makers

Information Governance Professionals

—

Central Government

Local Authorities

NHS

Transparency and Accountability

Greater public engagement and transparency surrounding health and social care data is required, addressing how data is gathered and shared, and the purposes for which it is being used.

A Government-led national communication campaign is needed to ensure the public is consulted, educated and informed about the use of their data. The campaign should set out the steps to be taken if people do not want their data to be used or shared in a certain way.

Information about data use, origin, and purpose should be communicated clearly to the public, including versions that are tailored to specific groups, such as young people, older people and people with disabilities.

An explanatory notice should accompany the publication of policy decisions, which makes clear the additional factors which are (by necessity) incorporated into statistical models and resulting policy decisions, including any limitations, uncertainties, and assumptions.

Any uncertainties, limitations or gaps in the data should be highlighted and the implications explained. Misunderstandings and false narratives should also be directly addressed and corrected where necessary.

Data Protection and Information Governance

Data acquisition, sharing, and processing should start and end with robust information governance. (By information governance, we refer to the controls and processes in place for the management of data pursuant to relevant legal obligations such as confidentiality and data protection.)

All data processing should undergo an impact assessment at the outset and be reviewed regularly. This should incorporate an ethical review and assessments of data protection, human rights and equalities, and algorithmic impact. Data ethics principles should be comprehensively defined, and consider potential impact on data subjects, risk of bias, and limitations in data.

Assessments should consider whether more privacy-preserving alternative approaches are available. Where possible, these assessments should be published.

For high risk or ethically controversial projects, the impact assessment should be reviewed by an independent, external data ethics committee, and the minutes of any meetings published.

Information governance controls, including details of limitations imposed, should be communicated to the public to provide assurance that risks have been considered and mitigated.

HEALTH AND SOCIAL CARE GUIDELINES

Target audience:

Policy and Decision-Makers

Data Analysts

—

Central Government

Local Authorities

NHS

Improving Data Quality and Robustness

Data quality and interoperability issues between systems have presented significant challenges during the pandemic, particularly when local and national responses intersect.

- Government investment in public data architecture is urgently required to address data quality and interoperability issues.
- Collecting, cleansing and harmonising data is an important step which must be provided with sufficient time and resourcing. Formal steps should be taken and recorded to ensure the independence and quality of raw data. Inaccuracies and omissions should be highlighted for rectification.
- As part of the National Data Strategy, a public sector data standards framework with clearly defined, consistent formatting, labels and ontology should be introduced and implemented to ensure interoperability and robustness of outputs.
- Qualitative data should also be collected and incorporated to support and contextualise quantitative findings (for example, to incorporate local knowledge or interview data).

Skills and Training

Additional training and resource is needed to ensure those tasked with making data-driven policy decisions are able to interpret, evaluate and interrogate statistical outputs and understand the associated risks and limitations involved when using data science and statistical modelling methods.

- Public sector decision-makers and policymakers should be required to undertake data literacy training, (incorporating relevant law and ethics issues) to ensure understanding of outputs and enable proper evaluation.
- Interdisciplinary working should be encouraged to help bridge gaps in understanding between policy makers, civil servants and social service organisations, law and governance specialists and mathematicians/data scientists.
- Professional standards for data science should be supported and implemented in the public sector. Ethical principles, linked to legal requirements and independent oversight, should form part of this professional framework.

HEALTH AND SOCIAL CARE GUIDELINES

Use of AI and Mitigating the Risks of Bias and Discrimination

Target audience:

Decision-Makers

Data Analysts

Information Governance Professionals

—

Central Government

Local Authorities

NHS

There are concerns that the use of AI, as well as certain digital and tech-driven pandemic responses (including: exposure notification apps and health status certificates) may have an unequal and disproportionate impact on those who are already vulnerable to discrimination, digital poverty and exclusion in society.

Where AI is employed, it must be subject to a prior independent review to establish that it is fit for purpose, based on a representative sample and based on unbiased data.

Those who build models must have an understanding of the impact that the particular model has, ensuring that those who are collecting data are considering who is represented in those data sets, as well as who is missing and for whom the data can reasonably be generalised. There is a need for external validation of models on a real-world, diverse population data set.

Public health emergencies may justify the deployment of an algorithm that makes (false positive) errors, applying a precautionary approach. However, attention must be paid to monitoring and evaluating the consequences of such an approach. It is also important to ensure transparency around the adoption of such an approach.

Any health status certification mandated by government should be subject to a strict oversight mechanism, which minimises the risk of bias and discrimination. Mandatory equality impact assessments should be carried out and efforts made to educate the public on the probabilistic, contextual and time-limited nature of certification, any applicable legal and regularity frameworks and any human rights concerns.

There is a heightened need for consultation and transparency to help people understand the potential benefits and weaknesses of different approaches to data collection and technologies, so that their potential value and limitations are not undermined by a lack of trust in, or low rates of adoption by the public.

LAW ENFORCEMENT GUIDELINES

Target audience:

Local Police Forces

NPCC

NPoCC

College of Policing

Transparency and Accountability

Transfer of health data to the police should be minimised and subject to strict independent safeguards. This should include transparency with the way that sensitive information is accessed, handled, used and retained.

- Where Memoranda of Understanding are agreed between health agencies and enforcement bodies regarding sensitive health data, it is vital that the terms and conditions are made as transparent to the public as possible.
- Transparency around the way that sensitive information is accessed, handled, used and retained should be the default position in similar situations in future.
- Where public health data is obtained by the police, its status in relation to other police databases should be made publicly explicit.
- Police forces should be proactive in publishing indicative data designed to improve transparency and public understanding of police work. Where necessary, this should come with appropriate caveats so as not to appear to overrule or duplicate official statistics.

Ameliorating 'Overload of Information' for Frontline Officers

Whilst there has been an increased 'thirst for information' regarding the police response to the pandemic, this has often been experienced as an 'overload of information' for frontline officers who are required to process constant updates from government, national policing bodies and senior force management.

- Clearer recording mechanisms at both police force level and national level are needed for the number and type of data-driven initiatives that police officers are committed to.
- Specific consideration should be given to the additional legal tools that forces would need in order to fulfil extra responsibilities effectively.
- Frequent consultation with officers to identify points of saturation would also be beneficial.

LAW ENFORCEMENT GUIDELINES

Methods for Improving Data Quality, Robustness and Handling Practices

Target audience:

Local Police Forces

NPCC

NPoCC

College of Policing

The quality, robustness and handling of data have been highlighted as key areas for improvement for policing stakeholders. Collaboration across all levels of policing is required to ensure that a consistent approach to these matters can be developed moving forward.

- Standardisation of labelling and terminology across key policing stakeholders is necessary to ensure that the quality of data entering data collection and analytics tools is of a high enough standard. This step is central for ensuring the robustness of outputs.

- National policing bodies should work with local forces to evaluate whether the technology developed or lessons learned in the context of the COVID-19 pandemic are applicable to other areas of policing business. This must, however, meet tests of proportionality to satisfy any concerns over mission creep.

- Enforcement of Fixed-Penalty Notices can have an impact on community-police dynamics. Having systems in place which scrutinise the frequency and distribution of tickets in comparison to other policing approaches is essential. The ability to triangulate indicative data with official data at regular intervals represents good practice and should continue.

- The development of data collection tools should be accompanied by a data protection impact assessment and follow the handling requirements in the Management of Police Information guidelines. This is particularly important when datasets are disclosed to third parties (such as academics).

- With some police forces using their employees' data in more innovative ways to aid planning and resourcing decisions, there should be more robust oversight of police handling of sensitive health data by the NPCC, with guidance from relevant bodies such as the ICO, to ensure a more consistent policy across the country.